Calhoun: The NPS Institutional Archive

DSpace Repository

Theses and Dissertations | 1. Thesis and Dissertation Collection, all items

1985

Proposed Data Administration strategy for the U.S. Coast Guard.

Jarrad, Thomas A.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/21510

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun

http://www.nps.edu/library

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

LIBRARY
DUATE SCHOOL
3

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

PROPOSED DATA ADMINISTRATION STRATEGY
FOR THE U.S. COAST GUARD

by

Thomas A. Jarrad

March 1985

Thesis Advisor: Norman R. Lyons

Approved for public release; distribution is unlimited

T222117

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)*: Proposed Data Administration Strategy for the U.S. Coast Guard	5. TYPE OF REPORT & PERIOD COVERED: Master's Thesis, March 1985
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Thomas A. Jarrad	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Postgraduate School, Monterey, CA 93943	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Postgraduate School, Monterey, CA 93943	12. REPORT DATE: March 1985
	13. NUMBER OF PAGES: 82
14. MONITORING AGENCY NAME & ADDRESS *(If different from Controlling Office)*	15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution is unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

data administration, data dictionary

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The purpose of this thesis is to provide the Coast Guard with an introduction to Data Administration (DA) concepts so that it may be better prepared to enter the fifth stage, the Data Administration stage, of Nolan's model of data processing growth. A brief history of data processing activities in the Coast Guard efforts related to DBMS's. Issues related to data dictionaries (DD's) and data dictionary/directory systems (DD/DS's) (Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

ABSTRACT (Continued)

are then presented including: the uses and benefits of DD's and DD/DS's and broad planning guidelines on how to implement a DD or DD/DS. The final two chapters are general recommendations to the Coast Guard on how to best prepare for data administration. These recommendations include developing: a central data dictionary, a DA charter, DA standards, and in-house training for general DA concepts and DBMS-specific topics.

Approved for public release, distribution unlimited.

Proposed Data Administration Strategy for the U.S. Coast Guard

by

Thomas A. Jarrad
Lieutenant Junior Grade, United States Coast Guard
B.S., Arizona State University, 1974
M.B.A., Arizona State University, 1976

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION SYSTEMS

from the

NAVAL POSTGRADUATE SCHOOL
March 1985

MONTEREY, CALIFORNIA 93943

ABSTRACT

The purpose of this thesis is to provide the Coast Guard with an introduction to Data Administration (DA) concepts so that it may be better prepared to enter the fifth stage, the Data Administration stage, of Nolan's model of data processing growth. A brief history of data processing activities in the Coast Guard is presented followed by an overview of current Coast Guard efforts related to DBMS's. Issues related to data dictionaries(DD's) and data dictionary/directory systems(DD/DS's) are then presented including: the uses and benefits of DD's and DD/DS's and broad planning guidelines on how to implement a DD or DD/DS. The final two chapters are general recommendations to the Coast Guard on how to best prepare for data administration. These recommendations include developing: a central data dictionary, a DA charter, DA standards and in-house training for general DA concepts and DBMS-specific topics.

NAVAL POSTGRADUATE
MONTEREY, CAL

TABLE OF CONTENTS

I. INTRODUCTION ---------- 7

A. PURPOSE OF THESIS ---------- 7

B. COAST GUARD ENTERING "DATA ADMINISTRATION" STAGE ---------- 7

II. HISTORY OF COAST GUARD DP ACTIVITIES ------- 10

A. FORMATION OF G-T ---------- 10

B. STANDARD TERMINAL CONTRACT ---------- 12

C. CG DATA PROCESSING TREND SETTERS -------- 13

1. Office of Command, Control and Communications (G-T) ---------- 13

2. 13th CG District Information Center-- 14

3. 12th CG District's ATONIS system----- 14

4. Honorable Mentions ---------- 16

III. CURRENT CG DBMS ACTIVITIES ---------- 18

A. DISTRICT DBMS EFFORTS ---------- 18

B. CG HQ SOFTWARE EVALUATION BOARD(SEB)----- 21

C. ANALYSIS OF ReQuest^{tm} DBMS ---------- 22

1. History of ReQuest ---------- 22

2. Product Overview of ReQuest --------- 22

3. Other Primary Users of ReQuest ------ 25

IV. ISSUES RELATED TO DATA DICTIONARIES --------- 26

A. WHAT IS A DATA DICTIONARY(DD)? ---------- 26

B. WHAT IS A DD/DS? ---------- 26

C. BENEFITS OF DD's AND DD/DS's ---------- 29

D. USES OF DD's AND DD/DS's ---------- 35

E. HOW MANY DICTIONARIES? ---------- 36

F. QUALITY OR QUANTITY IN A DATA DICTIONARY? 37

G. HOW TO IMPLEMENT A DATA DICTIONARY------- 40

1. Planning Dictionary Usage ---------- 40

2. Development of Dictionary Standards-- 42

3. Planning for Dictionary Integrity---- 48

4. Establishing Dictionary Security----- 49

5. Planning the Approach to Dictionary Creation ---------- 50

6. Selection of the First Application--- 50

V. ISSUES RELATED TO DATA ADMINISTRATION ------- 53

A. WHAT IS DATA ADMINISTRATION? ---------- 53

B. BENEFITS OF DATA ADMINISTRATION --------- 56

C. DATA ADMINISTRATION STANDARDS ---------- 59

D. DATA ADMINISTRATION CHARTER ---------- 63

VI. PROPOSED DATA ADMINISTRATION STRATEGY FOR U.S. COAST GUARD ---------- 69

A. DEVELOP CORPORATE DATA DICTIONARY ------- 69

B. DEVELOP DA CHARTER/STANDARDS ---------- 70

C. CONTINUE SOFTWARE EVALUATION BOARD------- 71

D. DEVELOP REQUEST IN-HOUSE USER TRAINING--- 71

VII. CONCLUSIONS ---------- 75

A. REMAIN FLEXIBLE ---------- 75

B. FOCUS ON THE END-USER ---------- 76

C. PLAN FOR DATA ADMINISTRATION -------- 77

LIST OF REFERENCES ---------- 79

INITIAL DISTRIBUTION LIST ---------- 81

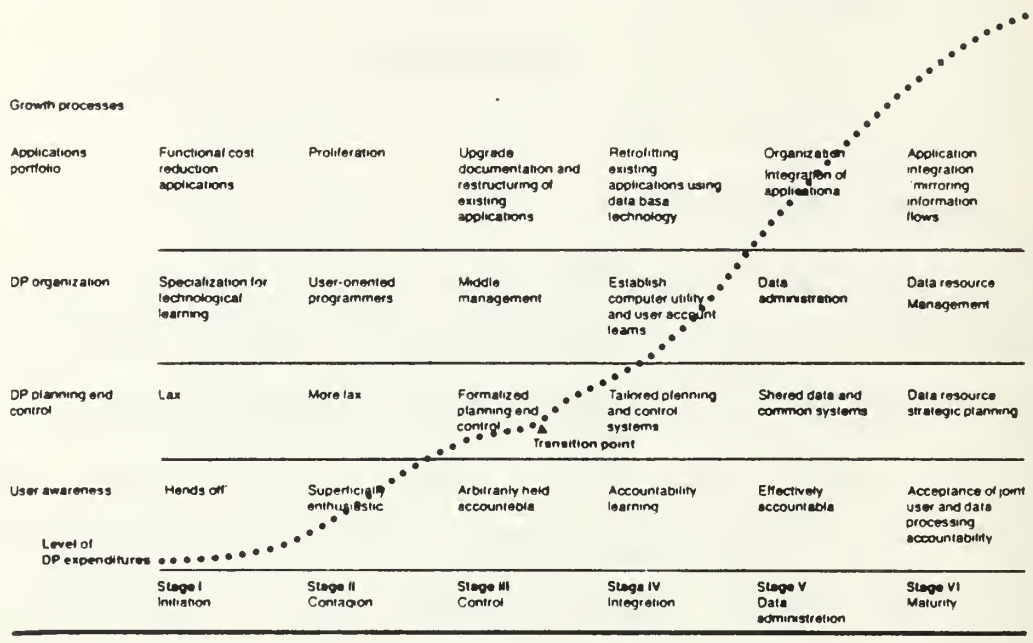
I. INTRODUCTION

A. PURPOSE OF THESIS

Originally, I had intended to do a market survey of commercial DBMS software packages and recommend one of these to the Coast Guard for use on their C3 minicomputer network. However, between the time I submitted my proposal and the actual writing of the thesis the Coast Guard Software Evaluation Board (SEB) selected the commercial DBMS, REQUEST^{tm}, as the software package it intends to support and encourage Coast Guard field units to purchase for the C3 minicomputers. Accordingly, the topic of my thesis shifted to a different theme, "A Proposed Data Administration Strategy for the U.S. Coast Guard." Now that the Coast Guard has made a commitment to a relational DBMS to meet its data processing needs it needs to devise a thoughtful and practical strategy on how to design and maintain the "data" that will be accessed by DBMS's throughout the Coast Guard.

B. COAST GUARD ENTERING "DATA ADMINISTRATION" STAGE

A widely accepted framework for understanding and evaluating data processing within organizations is Nolan's six stage model for the introduction and growth of the data processing function within organizations (see Figure 1-1).



Nolan's Six Stages of Data Processing Growth

FIGURE 1-1

Figure 1-1 graphically illustrates Nolan's six stage model. The rising dotted line represents the increasing level of expenditures in the total data processing budget of an organization. [Ref. 1: pp. 76-89]

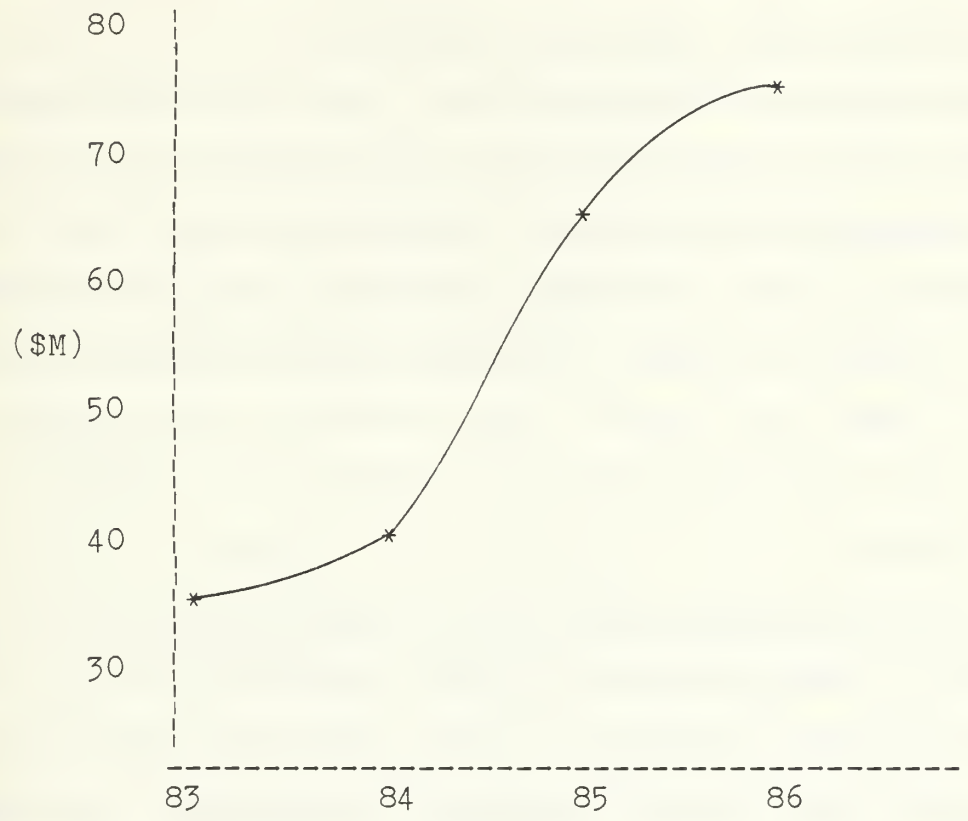
In general, I would place the Coast Guard in Stage IV, Integration. The Coast Guard is currently in the process of retro-fitting existing applications using DBMS technology. Clearly this is the Stage IV applications portfolio growth process as seen in Figure 1-1. It is important to note that data administration and data resource management are listed as the next two stages in the DP growth process. As the

Coast Guard enters the data administration stage it will need to shift its emphasis from managing hardware and software to managing "data". How to define and manage "data" will be the primary aim of this thesis. I will be discussing current topics related to data processing including: relational databases, data administration, data dictionaries, and data dictionary/directory systems (DD/DS's).

II. HISTORY OF COAST GUARD DP ACTIVITIES

A. FORMATION OF G-T

Prior to 1981 the Coast Guard had no formal structure in its organization chart for a data processing office. Most computing was centralized and performed by an Amdahl mainframe computer at the Department of Transportation's (the Coast Guard's parent department) Transportation Computer Center in Washington, DC. This mainframe is still being leased today to handle the Coast Guard's various accounting functions including paychecks to its approximately 35,000 civilian and military members. In addition to the centralized computing being done by the mainframe in Washington, DC the many operating units within the Coast Guard have also been making significant buys of microcomputers and word processors to handle their various local word and data processing needs. In fiscal years 1983-1985 the Coast Guard spent $36, $41, and $65 million dollars for local computing needs (includes hardware, software, supplies, services, and site preparation). The Coast Guard anticipates spending over $73 million in fiscal year 1986 for local computing needs. Figure 2-1 graphically illustrates the Coast Guard's increasing investment in data processing. [Ref. 2: p. 2-44]



Coast Guard DP Investment for FY's 1983-86

FIGURE 2-1

In March 1981 the Commandant of the Coast Guard formed a new office, the "Office of Command, Control and Communications (G-T)," at Coast Guard Headquarters in Washington, DC. The charter of this new office was: to establish Coast Guard wide data processing policies, standardize equipment and procedures where practical, and to establish a comprehensive and dynamic "Information Resources

Plan (IRP)". The Coast Guard intends to use its IRP as a roadmap to meet its data processing needs over the short range (3 years) and long range (10 years).

After G-T was established at Coast Guard Headquarters the 12 Coast Guard districts, under Headquarter's control and geographically spread throughout the U.S., also established "data processing divisions (dt)", within their district organizations. Both Headquarters and the districts staffed these new offices with personnel cannibalized from three other existing divisions: Electronics, Communications and Planning.

In 1981 the Commandant of the Coast Guard recognized a need for a formal structure within the Coast Guard to manage its increasing investment in data processing resources and implemented the new office relatively quickly. Initially, there was some resistance to the new office but now (G-T) and its district counterparts, the (dt) divisions, are well accepted and recognized as a vital part of the modern Coast Guard.

B. STANDARD TERMINAL CONTRACT

On 1 July 1981 the Coast Guard awarded a competitive contract to C3 corporation to purchase $40 million worth of minicomputers (Coast Guard Standard Terminals). These terminals are high-end micros to low-end minicomputers belonging to a compatible family of systems manufactured by

Convergent Technologies, Inc. which also manufactures a wide range of peripheral equipment (hard and floppy disks, printers, tape drives, modems, extended memory) and software packages and utilities. The CG Standard Terminal contract expires on 1 July 1986 and sets a maximum order limitation of 3,384 Standard Terminals (2,858 keyboard/displays and 526 cluster controllers) [Ref. 3: p. 13]. The Coast Guard has currently purchased over 3000 CG Standard Terminals [Ref. 4: p. 6].

C. CG DATA PROCESSING TREND SETTERS

1. Office of Command, Control and Communications (G-T)

In its official status of policy-maker, (G-T), has been involved in many projects that have benefitted the Coast Guard DP community at large. G-TPP, a division within G-T started an Office Automation project in 1981 that is testing the following Standard Terminal features in an integrated environment: Word Processing, Electronic Mail, Networking, Forms Editor, File Mangagement, Database Management, Multiplan, and Communications. This is an on-going project which will eventually generate a "Standard Terminal Office Automation Plan" which will help CG field units to take advantage of all the capabilities of the Standard Terminal. [Ref. 4: p. B-2]

2. 13th CG District Information Center

The 13th CG District in Seattle, WA, established an Information Center on 1 September 1982 with a charter to support end user computing. The initial thrust was to provide the maximum possible end user training so that each computer equipped unit would have a cadre of qualified operators. There is an on-going program of tutorial development and computer based training supplemented with some in-house classroom training. The Information Center is readily available for users to sit down with consultants to solve their problems and also to pursue computer based training on an individual basis. Enlisted personnel completing Information Center training are entitled to the appropriate qualification codes. The Information Center also maintains an extensive library of reference manuals and periodicals for walk-in use. The CG Standard Terminal Training Program at the Information Center includes the following courses: Computer Literacy, End User Introduction, Word Processing, Multiplan, Databases, Microrim/RBase 4000, Executive Orientation, System Manager Introduction, System Manager, Interactive Query Language, and Users Guide. [Ref. 5]

3. 12th CG District's ATONIS system

In 1984 the 12th CG District developed an "Aids to Navigation Information System (ATONIS)". ATONIS was a CG Headquarters sponsored project assigned to the 12th CG

District to replace a previous system "Semi Automated Navigation Data System (SANDS)". SANDS had been used by the Coast Guard to collect and analyze data on navigation aids throughout the Coast Guard between 1970 and 1983. Under SANDS, data was collected on complicated forms which were filled out by the field maintenance units, reviewed by the district office, and then keypunched by personnel in the finance department prior to being sent to Washington, DC. At periodic intervals (and whenever requested) output forms containing the latest data in the system were returned to the District Offices for verification and other uses. SANDS suffered from tedious data collection procedures, a high input error rate, and a slow information turn around time. It also did not collect all the data required by the district and field units. Despite numerous attempts to improve the system, in 1983 the system finally "collapsed". Collective protests by the District Offices over the high work load and low return led CG Headquarters to abandon SANDS and direct the individual districts to use locally developed systems designed for their own needs until a new national system could be developed. [Ref. 6: pp. 3-4]

The 12th CG District's ATONIS is the new national system for CG aids to navigation. The 12th District chose R:BASE 4000^{tm} to develop ATONIS. Using standard relational database concepts and practices they developed a standard data dictionary, standard relations (or tables), and a

series of standard reports. This new system is much more versatile and efficient than SANDS and should the need ever arise to transfer the data to a DBMS other than R:BASE 4000^{tm} the conversion should not be too difficult since relational database procedures are relatively standard across all DBMS software products (i.e. data dictionary, tables, reports, menus, command programs, etc.).

4. Honorable Mentions

Other CG Districts and HQ units deserve mention for their pioneering work in data processing within the Coast Guard. The 14th District has implemented a very fine semi-automated message handling system for message traffic within their geographic boundary. The 14th District also designed an automated system to monitor information on customers of the Coast Guard package store in Honolulu. This system was used to examine the buying patterns of the customers and then set a store policy to control those customers who were making excessive purchases. The 11th District implemented a comprehensive Search and Rescue (SAR) decision support system (DSS). Their DSS includes a graphics software package that produces the entire 11th district coastline along central California and key geographical points within that same area. With this system the 11th District can keep track of its ships visually on a computer terminal and respond to any distress calls with the ship nearest to the distress position. Finally, EELAB and EECEN are responsible

for most of the research, testing, and configuration control that is done with the CG Standard Terminal and its associated software and peripherals.

III. CURRENT CG DBMS ACTIVITIES

A. DISTRICT DBMS EFFORTS

Coast Guard District DBMS's currently in use were identified and compared in a recent report contracted to Electronic Data Systems Corporation (EDSC) by the 13th CG District [Ref. 7]. EDSC evaluated 9 DBMS products currently being used throughout the Coast Guard. These 9 DBMS products are:

ADS	Convergent Solutions, Inc. Ms. Darcy Kamp 118-35 Queens Boulevard, Suite 900 Forest Hills, New York 11375
ADEPT	Parameter Driven Software, Inc. 30800 Telegraph Road, Suite 382280 Birmingham, Michigan 48010
CT-DBMS	C-3 Inc. Mr. Bob Williams 11425 Isaac Newton Square South Reston, Virginia 22090
dBASE II	Ashton-Tate Mr. Jim Rowe 10150 West Jefferson Boulevard Culver City, California 90230
dBASE III	Ashton-Tate (same as above)
EMESIS	Electronics Engineering Laboratory LCDG Hugh Grant 7323 Telegraph Road Alexandria, Virginia 22090
IQL	C-3, Inc. (same as above)

R:BASE 4000	Microrim, Inc. Mr. Dennis Murphy 3380 146th Place SE Bellevue, Washington 98007
ReQuest	System Automation Corporation Ms. Laurie Livingston 8555 Sixteenth Street Silver Spring, Maryland 20910

The selected products were evaluated according to the following eight functions:

Data Manipulation. Capability for flexible data access. Responds to inquiries with speed and accuracy. Provides relational and mathematical operations. Modification capabilities include efficient updating of data and the database structure.

Report Capabilities. Capability to present information in flexible and user-defined formats.

Multiuser Capability. Capability that allows more than one user to be active in the same database. Deadlocking, file-locking and record-locking features are required to support a multiuser environment.

Data Integrity/Security. Capability to store and protect information from unauthorized users. Controls data access and prevents input or revision of unqualified information.

Manufacturer Support. Willingness of the manufacturer to respond to the product survey letter, to aid in the development of applications and diagnosis of problems, to provide user training, and to plan for future products.

Ease of Use. Ability for the average USCG user to install, learn and make effective use of this product.

Specifications. Requirements of hardware and software to support normal product application. Ability to function in the USCG Standard Terminal environment.

Compatibility/Portability. Capability for communications with other frequently used software products, including input and output of data sets in acceptable formats. Provide for telecommunications and allow access by any user program. [Ref. 7: pp. 12-13]

Using the critical functions listed above, EDSC was able to narrow the evaluation down to 2 DBMS's: R:BASE 4000tm and ReQuesttm. After re-examining these two options EDS Corp. selected R:BASE 4000 as the better product.

B. CG HQ SOFTWARE EVALUATION BOARD (SEB)

On September 20-21 1984 a Software Evaluation Board (SEB) was held at CG Headquarters to determine what DBMS software package should be recommended as a CG "standard" DBMS. The intent of the SEB was to select a commercial DBMS and then advise all CG field units that this particular DBMS would be supported by CG HQ via users guides, training programs, documentation, and in some cases funds to purchase the software. This approach encourages the users to voluntarily use the selected "standard" DBMS but still leaves them with the freedom to use other software packages if they so desire.

A memo was sent out on 30 August 1984 to all G-T and District (dt) divisions interested in selecting a standard DBMS for the Coast Guard. 22 people responded to the memo and were invited to attend the SEB. Out of the 22 only 9 people were selected as members of the SEB. However, the other 13 people were there to offer their opinions and present papers supporting their view on the best DBMS for the Coast Guard. The 9 members of the SEB were an executive committee of G-TT, G-TDS, and G-TPP representatives, who ultimately made the decision.

The SEB selected ReQuest^{tm} as the current Coast Guard "standard" DBMS. They made this decision after reading the 13th CG District's report on an evaluation of 9 different DBMS software packages [Ref. 7], the papers sent to the

board from the various districts supporting different DBMS's, and in particular, after hearing the opinions of all the people attending the 2 day meeting. The recommendation by EDS to select R:BASE 4000 was only one of many factors considered by the board before making its decision to select ReQuest as the standard CG DBMS. Both products are good but the SEB felt ReQuest was just a little bit better.

C. ANALYSIS OF REQUEST^{tm} DBMS

1. History of ReQuest^{tm}

ReQuest was initially developed in the early 1970's for mainframe data base programming. It was widely used in the Army and the airline industry before it was redesigned to function at the micro level and released for sale in November 1983. It is designed to run in a multi-vendor environment under the MS-DOS, PC-DOS, and CTOS operating systems.

ReQuest automatically converts mainframe data formats into ReQuest data base formats and permits multiple access simultaneously for report generation. It can maintain directories of reports and forms created during system use. [Ref. 8: p. 46]

2. Product Overview of ReQuest^{tm}

Overview. ReQuest is manufactured by System Automation Corporation (SAC). The ReQuest system has been

divided into 5 main modules: search/report, data entry, data dictionary, menu maintenance, and security. ReQuest is menu-driven and contains no procedural language.

Data Manipulation. ReQuest allows a full scope of data manipulation functions, including automatic computational options and quick, full-range retrieval functions. ReQuest's uses a B-Tree search function. ReQuest supports the relational operators select, join, and project.

Report Capabilities. The ReQuest report mode is very flexible and is capable of representing the selected data in graph format. ReQuest does not allow total free formatting of reports, yet provides for much customization.

Multiuser Capability. ReQuest can support a multiuser environment. The use of ReQuest security levels prevents specific users from accessing the same data through a deadlocking prevention system.

Data Integrity/Security. ReQuest does not have an integral data recovery program. ReQuest allows users to be assigned a security level and password. The security levels range from 1 to 9.

Manufacturer Support. ReQuest has some manufacturer support for their product. ReQuest representatives are readily available to answer technical questions by telephone. ReQuest maintains a user "hotline" for giving technical advice. User training is provided upon purchase of their product. ReQuest is also delivered with a manufacturer-provided tutorial. However, this does not mean additional training programs should not be set up to support this product. In my opinion an in-house training program is almost a necessity since the ReQuest tutorial and reference manual are a little beyond the understanding of the average end-user.

Ease of Use. ReQuest is a menu-driven system which makes it relatively easy for the user to create a common application. The ReQuest tutorial is informative and serves as a useful learning tool.

Specifications. ReQuest is fully-compatible with the CG Standard Terminal hardware and software.

Compatibility/Portability. ReQuest is accessible for teleprocessing with the use of CT-NET, which supports ISAM. ReQuest is available for all CT hardware, MS DOS, and HP150 PC. [Ref. 7: pp. 28-30].

3. Other Primary Users of ReQuest^{tm}

C3 Incorporated was awarded a $73 million contract on September 1984 by the U.S. General Services Administration (GSA) for office automation systems. The contract award is through September 1985, with eight fiscal year renewal options. The contract calls for C3 to provide GSA with systems equipment from Convergent Technologies (CT), Inc. C3 will provide up to 4,299 CT workstations, including installation, software, training and system maintenance for various GSA offices nationwide. ReQuest^{tm} DBMS will be part of the standard software provided with each system. Certainly the news of this contract must have had an impact on the selection made by the CG Software Evaluation Board. [Ref. 9: p. 87]

IV. ISSUES RELATED TO DATA DICTIONARIES

A. WHAT IS A DATA DICTIONARY?

A data dictionary is a mechanism to collect, maintain, and publish information about data. It is a central repository of metadata (information about data). Basically, a data dictionary provides a mechanism to define and use information about data elements, groups of elements (records or segments), groups of records (files or databases), and the relationships between these entities. It is also capable of defining other entities, such as input forms, reports, screens, processes, procedures, and just about anything else. All data definition entities are built on the foundation of the element definition. [Ref. 10: p. 1]

B. WHAT IS A DD/DS?

Data dictionaries are often identified as data dictionary/directory systems (DD/DS). These systems are capable of not only storing metadata , but are also capable of providing cross-reference information (directory). The dictionary provides information about what the data is and how it is used. Thus, the dictionary provides a logical view of the data while the directory provides information on where the data physically resides and how it can be accessed. [Ref. 10: p. 3]

All DD/DS's provide the basic functions necessary to capture and maintain metadata and to generate reports from that store of metadata. Some DD/DS's also have the ability to generate data descriptions and program code and to support test environments. Data descriptions are often taken from an existing DBMS and loaded directly into the dictionary. [Ref. 11: p. 181]

Two types of reports are provided by DD/DS's: dictionary listings and cross-reference reports. The dictionary listings list all the data entries alphabetically or by entry type. In the cross-reference report data entries in the dictionary are associated by the relationships in which they participate. Since these relationships are bi-directional, the cross-reference may be either top-down or bottom-up. For example, one may ask to see a top-down listing of entries associated with a particular application system. One could also ask for a trace of all entries with which a particular element is associated, a bottom-up view. Some selectivity may be used with regard to the entries displayed. For example, one may wish to see only those programs associated with an appication system, not databases or elements. Selectivity may also be applied to the scope of information displayed for each entry. For example, one may wish to see only the names of those entries associated with element X, not the full information on each. [Ref. 11: p. 182]

Most DD/DS's provide a selection of preset reports that can be executed by the user directly. Some also provide a query language so that users may formulate reports of their own choosing. If the dictionary data base is maintained in a standard DBMS format, the reporting features are normally extended to include the report generator or query language facility available with that DBMS. [Ref. 11: p. 183]

The directory function of a DD/DS makes it the point of contact between application programs and the database. In such environments it is valuable to be able to define a number of statuses or conditions under which the objects defined will be used. For example, if a file is being modified, the directory should reference the old version of the file until changes are complete and have been verified. Then the new version of the file should be referenced. If the DD/DS does not allow differences in status, e.g., old and new, the two definitions cannot exist simultaneously. [Ref. 11: p. 183]

In this thesis I use the terms data dictionary and DD/DS synonomously with the understanding that these two tools have significant differences in capabilities and use. The data dictionary is more of a passive tool and the DD/DS is more often used in an interactive environment.

C. BENEFITS OF DATA DICTIONARIES AND DD/DS's

Although there has been an increase in the efficiency of methods used to collect, compute, and distribute data, there is still a void in the understanding of the characteristics and relationships of the data itself. It would be unreasonable to expect an engineer or contractor to construct a high-quality building without understanding the characteristics of the building materials. Yet data processors often attempt to build high-quality systems while ignoring the characteristics of the raw material of data. There should be an interest in defining and documenting information about this raw material. The data dictionary is a tool for the effective utilization of data. It enables use to use data effectively, efficiently, and consistently. [Ref. 10: p. 18] The benefits of data dictionaries and DD/DS's include the following:

1. Enhance corporate survivability---Data compiled about a company is an important corporate asset. Accurate information about how a company functions and about its employees and clients is vital to the success of any corporation. It is not difficult to measure the value of such data. According to the findings of a recent survey, only two out of ten companies whose data centers were destroyed were still in existence one year after the catastrophe. Any sensible data center will take great precautions to safeguard the company's data including its data dictionary to insure its survival. [Ref. 10: p. 18]

2. Promote egoless knowledge---Structured programming techniques are accepted today as the best way to write programs. The concept is to code a program in such a way as to make the logic path easy to follow and easy to read. Programs are more easily understood by many people, thus easier to maintain. Programs are no longer the private property of a single author, because the logic is shared by other programmers. The code becomes the public property of the entire programming staff. For this reason, structured programs were sometimes referred to as "egoless" programs. A valid comparison can be made between structured programs and a data dictionary. The knowledge that a programmer or analyst has about a company's data and systems should be accessible to and shared by all members of the organization via the data dictionary. Employees are paid to become proficient in the knowledge of the company and they should share this knowledge with everyone in the company. Knowledge should be the public property of the corporation. If structured programs are egoless then the data dictionary represent egoless knowledge. Data dictionaries also save a considerable amount of time being spent in question-and-answer sessions. Instead of going to the experienced people on the data processing staff with questions end-users could go directly to the data dictionary. [Ref. 10: pp. 18-19]

3. Improve corporate communications---The data dictionary is a central repository of information that can

be accessed by all areas of a company. Unlike the traditional means of communicating by memo or other paper formats, the data dictionary is not limited to a specific distribution list. Anyone with a terminal and knowledge of the dictionary has access to all of its information. Another use of the data dictionary is as a glossary of terms. Many employees would not consider their office to be complete without a Webster's dictionary. In many ways, the data dictionary is like Webster's dictionary; it contains a glossary of terms used by the firm. The data dictionary is essential for clear communication within the corporation. As a glossary of terms, the data dictionary can be an invaluable education tool for new employees in data processing and user areas. [Ref. 10: pp. 19-21]

4. Support structured system analysis and design---An interactive DD/DS can be a very effective tool to support structured analysis and design. It can be used to document data store, data flow, and process entity types. As such, it is an efficient way of portraying system design details to the user. It can also be used to generate file, segment, and record definitions for a variety of programming languages. By doing so, we can centralize the control of program data definitions. This will ensure consistency of data use and inhibit data redundancy. [Ref. 10: p. 21]

Because we can centralize control of data use, the data dictionary can be a very effective tool in change-control

management. Since the data dictionary is the origin of all data definitions, any new data requirements must have the knowledge and approval of data administration. Because the dictionary enforces consistency of data naming and format, it significantly reduces the cost of program maintenance. System maintenance requests involving the expansion of data elements like payroll numbers, account numbers, and zip codes are good examples. In a system using a nine-digit zip code, it would be possible to identify every occurrence of ZIP-CODE prior to implementing the change and estimate the costs involved in making that change in the data design. [Ref. 10: p. 21]

5. Provides a better medium for system documentation--- The interactive DD/DS is superior to the word-processor as a documentation medium. Paper is more likely to be damaged than the magnetic medium of a dictionary. Documentation in a DD/DS is available to anyone who has access to a computer terminal. Documentation in an interactive DD/DS is "living," perpetual documentation; documentation on paper has a limited life span. Although it is possible to perform automated searches on word processing equipment, such searches are often limited to a single document. The automated search and cross-referencing tools of a DD/DS can span multiple programs, systems, databases, report definitions, form definitions, and other categories of documentation. [Ref. 10: p. 22]

The interactive DD/DS can greatly improve the reliability of documentation about the data used in a system. The DD/DS contains the documentation about the data definitions. By using the DD/DS to generate the source program data definitions, the data portion of the program is actually derived from the documentation. This direct link between documentation and system definition guarantees the accuracy of the data documentation. After a system is implemented, all data changes should first be made in the dictionary. Then the source code data changes can be generated from the DD/DS. This assures that the data documentation will be kept up-to-date, and the data processing staff will have more confidence in its accuracy. [Ref. 10: pp. 22-23]

6. Generates data definitions automatically---A major benefit of the DD/DS is its ability to generate data definitions for a variety of software languages. Some DD/DS's can generate file and record layouts for use in application languages such as COBOL and PL/1. Some DD/DS's can also provide data definitions for procedureless query or report languages such as NOMAD and FOCUS. Some DD/DS's can also automatically provide data definitions for several database management systems such as ADABAS and IDMS. Several major software vendors have combined their dictionary and DBMS products so the DBMS schema and subschema definitions can only be produced from the

dictionary (these are called "dependent" data dictionaries). Some commercial DD/DS's are also capable of generating COBOL Procedure Division source code from macroinstruction statements contained in the dictionary. [Ref. 10: p. 23]

7. Increases end-user involvement---Mini's, micro's, fourth-generation languages, and user-friendly inquiry and reporting tools are means by which we can utilize the user in the development of data processing systems. The data dictionary is one more tool to increase user involvement in system development. In the traditional development scheme, the user is only a reviewer and auditor of the system development efforts. The user really does not actively participate in the analysis and design effort itself. The data dictionary, when used with other modern development aids, can help balance the DP-user staff workload. The data dictionary is a tool to more effectively utilize the talents of both user and data processing personnel. By defining the characteristics, relationships, and editing criteria of the data, the user can be directly involved in the design of the system. The user will have more direct control over system design and, at the same time, save data processing staff time in the definition of data. The data dictionary is a tool to delegate more of the data processing workload and responsibility to the user community. [Ref. 10: pp. 28-30]

D. USES OF DATA DICTIONARIES AND DD/DS's

Data dictionary information falls into two categories: process entities and data entities. Process entities include such items as systems, programs, modules, and submodules. Data entities include files, database schemas, subschemas, records, segments, groups, and data elements. The data dictionaries and DD/DS's can be used to answer the following type of questions:

1. What programs are in system X?
2. What subroutines are called by program X?
3. Subroutine X is called by which programs?
4. Data element X is used in which records?
5. Which programs use record X?

Data dictionaries are also used for the following functions:

1. To store entities used in existing production systems and entities created during new application development.
2. To store proprietary and nonproprietary entities.
3. To document and control the procedures involved in the creation and evolution of these entities.
4. To record and monitor the events during the life cycle of a new application development.
5. To manage the tasks involved in data modeling and logical database design.
6. To provide change control of entities.

7. To store metadata and data.

8. To document standards, policies, and procedures. [Ref. 10: pp. 130-131]

E. HOW MANY DICTIONARIES?

When implementing a data administration function within an organization, the data administrator must address the following question: How many dictionaries should be developed to support the information resource management needs of the company? For a small company with only one location, the answer to this question is obvious. But for a large organization with many locations and many divisions the answer is not so apparent. For large organizations with several divisions the number of data dictionaries implemented depends upon the commonality of data used by the various areas of the company. The data administrator must research the degree of commonality by answering questions such as the following:

- Which data elements or entity classes are common to the different areas of the organization?

- What does the personnel division have in common with the accounting division?

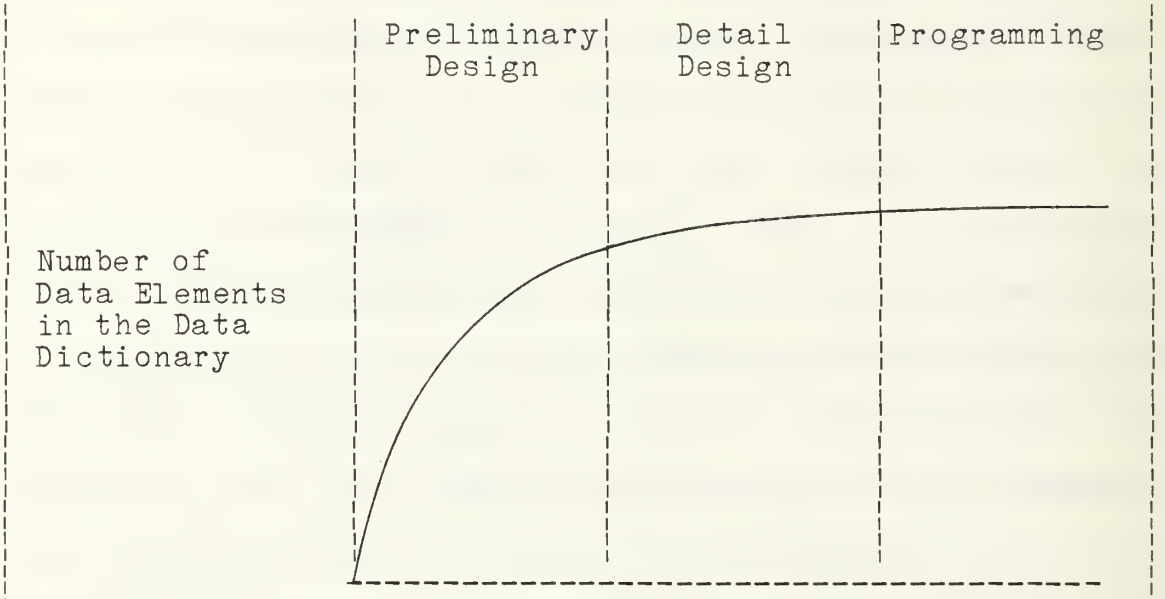
- What data elements are shared by both the electronics division and the R&D division?

- What information is common to both CONUS (continental U.S.) and overseas divisions? [Ref. 10: p. 133]

Although large organizations seem disjoint, there is often a significant amount of data common to all areas of the company. Personnel, budget, and accounting are examples of entity classes that are often shared by the entire organization. Even though an organization may have divisions that are spread over large geographic area, their information resource needs could be satisfied with one central dictionary. This is accomplished by downloading segments of the central dictionary to remote locations within the organization. This will provide metadata to a multitude of remote-site dictionary users. However, any updates or changes requested by the end-users should be channeled through the data administrator. [Ref. 10: p. 134]

F. QUALITY OR QUANTITY IN THE DATA DICTIONARY?

Figure 4-1 illustrates the evolution of data elements during the development of a new data processing system. During the preliminary design, all data elements in the existing user views should be identified and stored in the data dictionary. Figure 4-1 also illustrates the dramatic increase in data dictionary items during the preliminary design. The number of data elements defined during the preliminary design should represent approximately 80 percent of the data elements in the final implemented system. [Ref. 10: p.156]



Evolution of Data Elements During Development of New System

FIGURE 4-1

During the detail design, there should be a much smaller rate of increase in the number of data elements added to the dictionary. During this phase, data administration will add any new data elements to support future or anticipated user views. Other, additional data elements will be those concerned with system and program operating controls, auditing, and entities. These data elements include program-to-program controls, counts of data element values, counts of the number of data elements, or the number of

records and segments moved or transmitted. Once the programming effort has begun, there should be very few data elements added to the design of the system. Of course, there will be a few new data elements as a result of omissions in data design during the detail design phase. The further into the development of a project, the more closely the data administrator should scrutinize additional new data elements. For each new entity during the latter stages of development, the data administrator should ask:

---Is this new data element a duplication or variation of a data element that already exists?

---Why was this data element not introduced earlier in the design? Has there been a design change to justify the need for this new data element? If so, what impact will this new data element have upon existing data elements. Has this design change been approved by management? [Ref. 10: pp. 156-157]

Figure 4-1 presents an important principle of the data dictionary population during the life cycle of a new system. During the latter stages of the project, the quantity of data elements can be directly related to the lack of quality of the data design. A steady increase in the number of data elements during the detail design and programming phases might be an indication of incomplete data design during the preliminary design phase. [Ref. 10: p. 157]

G. HOW TO IMPLEMENT A DATA DICTIONARY

Successful data dictionary implementation is achieved through usage planning, procedure development, and the adoption and enforcement of standards and conventions for a variety of dictionary functions. One methodology for data dictionary implementation is composed of six stages: [Ref. 12: p. 1]

1) Planning dictionary usage
2) Development of dictionary standards
3) Planning for dictionary integrity
4) Establishing dictionary security
5) Planning the approach to dictionary creation
6) Selection of first application

The following section will look at these six strategic factors in greater detail and identify many of the considerations that should be taken into account.

1. Planning Dictionary Usage

A data dictionary does not bring benefits as an automatic result of its existence. It requires careful planning and directed effort to achieve gains. The data administrator is responsible for planning the usage of the dictionary. The first step in the plan is to identify the potential users of the dictionary. The users are either: corporate users or EDP users. EDP users include: systems development, systems maintenance, and operations personnel. Corporate users include business analysts, auditors, and

those departments served by EDP. Planning for the corporate users involves the following activities: [Ref. 11: p. 12]

a) Establish procedures to determine which department is the ultimate owner of a particular data entity.

b) Define the extent of a corporate user's involvement in dictionary information.

c) Agree on the extent to which a user might make use of data dictionary commands.

d) Institute procedures for liaison with the data administrator, and for the regular reporting of any additional users of a data entity.

Taken together, these operating guidelines have the multiple effect of easing dictionary development; bringing the corporate user into a closer relationship with EDP; and providing the means to develop that relationship. This is accomplished through the dictionary commands that enable users to carry out their own impact of change analysis without having to submit such requests through EDP. [Ref. 12: pp. 12-13]

Planning for the systems development personnel includes the early establishment of ground rules to:

a) Clearly define the controlling role of the data administrator's office.

b) Define procedures for the allocation of test views of data to each development team.

c) Introduce tight regulatory controls over the change from test to production view particularly where this involves changes to existing definitions.

Planning for the systems maintenance personnel should include:

a) The provision of full dictionary interrogation facilities.

b) A procedure whereby maintenance may request a new entity via the data administrator's office.

c) The potential supply of a test view dictionary for the maintenance group to use as a "scratch pad." This should be stringently protected to prevent potential corruption of production data definitions.

Finally, operations personnel would be using the data dictionary to obtain job set-up instructions and as a management aid for the administration of mass storage. They should be provided with:

a) Access to information such as physical file attributes and where those attributes are used.

b) Job stream components and interrelationships.

c) Possible update facilities for entities representing disk packs. These would be defined in terms of the data sets held on those entities. [Ref. 11: pp. 12-13]

2. Development of Dictionary Standards

Perhaps the most important aspect in developing a data dictionary is adopting the standards that will guide

its use. Without standards the dictionary will only automate and continue any existing chaos. The standards that should be addressed are:

a) Naming conventions for primary names of data and process entities.

b) Naming conventions for index or catalog names.

c) Standards for entity definitions.

d) Standardized data collection forms and procedures.

The standard which users usually identify as offering the most immediate benefits is the one related to the names of entities. Various methods of standardization for entities have been used and include the following:

- Coded names
- Titles
- Program names
- "OF" language
- Meaningful abbreviations
- Abbreviation by removing vowels

There is no one technique better than another. The important thing is to apply one standard consistently to all objects in the data dictionary. This is necessary so that data redundancy can be reduced; so that retrieval of data dictionary information can be performed in a coherent manner; so that data and processes can be recognized and distinguished; and so that some sort of understanding of the

data entities can be determined from their names. [Ref. 12: p. 13]

Objects that are to be defined in the data dictionary fall into three categories:

a) Physical Objects---Objects that are readily identified by an external unique identifier, that is in widespread usage outside the dictionary.

b) Logical Objects---Objects that or logical or conceptual in nature such as data elements, dataflows, processes, and functions.

c) Local Objects---Objects unique to a specific programming language which generate record descriptions in a data dictionary. [Ref. 13: p. 2]

There seems to be no better way to determine the purpose or meaning of an object than to require the user to write a precise narrative containing all the pertinent facts concerning the object in the real world. If the object is eventually added to the data dictionary, then this narrative should become an essential part of the dictionary definition (e.g., the description). [Ref. 13: p. 2]

Before adding an object to the dictionary the user usually wants to check to see if that object is already listed. Using a dictionary name to find a pre-existing object is not very reliable. Perhaps the best technique is to classify each object when it is added to the dictionary with a set of KWOC (Key Word Out of Context) values. Pre-

existing objects are then found by searching the dictionary for definitions having a matching set of KWOC values. This type of search eliminates name length restrictions and word sequencing, two of the four factors that make name searches unreliable. [Ref. 13: p. 2]

A useful validation of object definitions for completeness is to require that it contain at least one "prime word", and only one "class word". Of course it is necessary to compile a list of such words for each organization. Prime words will be industry related and therefore will differ for each organization. For instance, TRACEN, RADSTA, SUPCEN, WHEC, OFFICER, and ENLISTED, are of prime importance to the Coast Guard because they collect facts or data about them. Class words categorize different types and representations of data and therefore tend to be universal. [Ref. 13: pp. 2-3]

IBM created a technique for forming unique, readable data object names called the "OF" language. In the OF language an object name is composed from one class word followed by one or more "modifier" words. The class and modifier words are separated by one of several different "connectors". Class words in the OF language are identical in scope to the previous paragraph. The most frequently used class words are: NUMBER, NAME, TEXT, CODE, QUANTITY, DATE, AMOUNT and FLAG. The symbols: # , N, T, C, Q, D, $ and F respectively are used to denote these class words. The connectors are:

space = OF

* = WHICH IS or WHICH ARE

: = OR

& = AND

- = used in hyphenated words

/ = BY or PER or WITHIN

Table 4-1 presents some "OF" language descriptions and descriptors:

TABLE 4-1

"OF" Language Example

OF Language Description	OF Language Descriptor
NUMBER OF EMPLOYEE	# EMPLOYEE
NAME OF EMPLOYEE	N EMPLOYEE
NAME OF EMPLOYEE WHICH IS LAST	N EMPLOYEE*LAST
CODE OF EDUCATION-- LEVEL OF EMPLOYEE	C EDUCATION--LEVEL EMPLOYEE
NUMBER OF DEPARTMENT OF EMPLOYEE	# DEPARTMENT EMPLOYEE
AMOUNT OF RAISE	$ RAISE
AMOUNT OF SALARY WITHIN MAXIMUM	$ SALARY/MAXIMUM

The definitions in Table 4-1 are derived by successively modifying the appropriate class word, starting with the most significant modifier, then the next most

significant and so on. An OF language descriptor is formed by substituting the various symbols above into the definition. [Ref. 13: p. 3]

There does not seem to be any easily remembered formula that will return acceptable abbreviations in all cases. The most effective way to abbreviate is to translate a word to its acceptable abbreviation by looking it up in a "standard word and abbreviation glossary." Users who have automated the process of producing standard abbreviated names seem to agree on the following guidelines:

a) Each organization should develop a standard glossary containing all words approved for use when generating dictionary names.

b) If an attempt is made to use a word in a name and that word is not in the approved standard glossary, then a decision should be made either to add it to the list or to not use the word.

c) Each word in the standard glossary must be given one acceptable abbreviation.

d) It is useful to indicate that a particular word, if used to generate a name, should always be used as the first or last part of the name, or should be dropped completely from the name.

e) Marking prime and class words in the list enables a validation for completeness to be performed on the chosen object identification, such as with the KWOC values.

f) As the standard abbreviations become widely known, it is beneficial to permit a word and its abbreviation to be used interchangeably such as when KWOC searches are performed. [Ref. 13: p. 4]

3. Planning for Dictionary Integrity

Dictionary integrity means insuring that the data loaded into the dictionary is correct, and remains correct. It is essential to the future development of the data dictionary, and to the quick acceptance by users of data resource concepts, that users have the utmost confidence in the data dictionary system. This confidence can only be gained if the data in the dictionary is accurate and reliable. Planning for integrity involves the auditing and validating of all matters relating to input, output, and update of the dictionary. Some of the things that should be considered for data integrity include:

a) Are there sufficient administrative checks that take place prior to a member amendment; or the upgrading of a member's test view to production status?

b) Are there adequate administrative or computerized checks for the enforcement and validation of standards?

c) Have adequate procedures been designed for back-tracking and reporting on any violation of standard or convention?

d) Do other programs access the dictionary directly?

e) Are the dictionary system's command and query languages provenly effective and "safe"?

f) If operating in on-line mode, what are the checks against concurrent updating?

g) Are the error recovery facilities comprehensive or do they require additional support from EDP? [Ref. 13: p. 11]

4. Establishing Dictionary Security

When planning data dictionary security it is easy to get carried away and forget that the system has to be usable as well. Not only must it be usable from outside, but the security provisions imposed should not be so complex that the data administrator is forced to spend an inordinate amount of time controlling it. In general there are three main topics to be considered in dictionary security: physical safety, access control, and external requirements (such as those imposed by auditors). With regard to physical security normal back-up copies should be made and copies of each transaction should be kept in the event of a breakdown. As a general rule it is wise to copy everything, via automatic transaction logging, and to keep the copies safely off-site. Access security and external requirements can be imposed by the DBMS that the data dictionary is operating under. The protective mechanisms used by the DBMS

include: passwords, data encryption algorithms, and restrictions on the availability of access and update commands. [Ref. 14: p. 11]

5. Planning the Approach to Dictionary Creation

There are basically two ways an organization approaches creating a data dictionary: top-down or bottom-up. The top-down approach makes it possible to ease into dictionary usage, and provides a step-by-step learning process from least to most complex. It also maximizes the usefulness of the dictionary at a high level from early on, thus "spreading the word" faster and more effectively than would be possible with any other approach. The top-down approach also prevents the problem of synonyms or homonyms appearing in the dictionary, because each downward step is uniquely defined before loading. [Ref. 14: pp. 11-12]

The bottom-up approach makes use of the corporate glossary that contains, at the data element level, an absolute or "pure" definition of every data item used by an organization. This "pure" base is then used as the reference point for all development and maintenance in the future, and for rationalizing the chaos of the past. It is a desirable goal but difficult to achieve because of the large volumes of data entities involved. Most organizations end up adopting a mixture of the two approaches discussed above. For example, the top-down approach might be used for system development projects and the bottom-up for

maintenance projects. [Ref. 14: p. 12]

6. Selection of First Application

The choice of the first application to be defined in the data dictionary is one of the most strategic decisions the data administrator will make regarding dictionary implementation in an organization. This choice has been found, on many occasions, to be the key to the eventual success or failure of the project, and the decision therefore justifies a considerable amount of time being spent on it. Uppermost in the mind of the data administrator should be the need to balance visible achievement against longer term aims of management and control. [Ref. 14: p. 12]

A good approach is to take a specially selected existing system and to analyze that to implement the dictionary. Every organization has a system which is small, neat, and apparently self-supporting. This is the ideal place to start. It can be cost justified, and the volumes are small enough to manage through the crucial growing pains of dictionary usage experience. If a small task force can also use the dictionary for development work, then a combination of visible success and hidden achievement can be accomplished while building knowledge for the users, and definitions for the dictionary. This approach achieves two objectives: [Ref. 14: p. 12]

a) It wins the crucial support of the users who are able to see a system implemented and working and allows them to see the benefits of the data dictionary quickly.

b) It gives the data administrator some hope for upper level management and user support for starting a major project like constructing a corporate glossary and eventually a corporate data dictionary.

V. ISSUES RELATED TO DATA ADMINISTRATION

A. WHAT IS DATA ADMINISTRATION?

If you ask a programmer or analyst what is data administration he will most likely say it has something to do with data dictionaries or databases. While this answer is not incorrect, it merely describes some of the tools or facilities used by the data administrator. These tools are only a means to the overall objective of the data administrator (DA) which is to plan, document, manage, and control the information resources of the entire organization. Data dictionaries, DD/DS's and databases help us achieve this goal, but none are an end in themselves. The role of the DA is not to maintain individual databases and dictionaries but rather to integrate and manage corporation-wide information resources by "using" data dictionaries and well-designed data structures. [Ref. 10: p. 3]

To maximize the return on investment from a data dictionary, the DA must provide management with answers to the following questions:

1) What will be achieved by implementing a data dictionary? What are the costs and benefits associated with its implementation?

2) What information should be loaded into the data dictionary?

3) Who will be responsible for inputting information into the dictionary?

4) Who must review and approve this information before it is entered in the dictionary?

5) What steps will be taken to insure the quality of information before it is entered?

6) Once information is loaded into the dictionary, how will it be maintained?

7) Who is responsible for maintaining the integrity of the data in the dictionary?

8) Will the dictionary be used for developing new systems or for assistance in maintaining existing systems?

9) Which software languages should be supported by the dictionary?

10) What DBMS's should the dictionary support?

11) Will this dictionary be used by the entire organization, by individual departments, or individual application development projects? Should the DD be used to document data or process definitions, or both?

12) Who will be the end users of the dictionary?

13) What training will be necessary for users of the dictionary?

14) What should be the first project or application using the dictionary?

15) In what sequence should other projects or applications be added?

16) What are the short- and long-term objectives of using the data dictionary?

17) Does management understand all of the capabilities and facilities of the dictionary? [Ref. 10: p. 4]

The objective of the DA should be to answer these type of questions before the data dictionary is implemented. By doing so, an organization can assure itself that the implementation of a dictionary will be sensible and cost-effective.

The role of data administrator (DA) is often confused with that of the database administrator (DBA). The difference between these two positions is significant and should be noted. Normally, DBA's are responsible only for the design, implementation, security, and maintenance of physical databases. It is the responsibility of DA's to determine the contents and boundaries of each database. The DA first builds a logical model of the database which is later implemented by the database administrator (DBA). This is analogous to the distinction between a systems analyst and a systems designer. Before the DA and DBA design a single logical and physical database, the DA should strive to plan and coordinate the construction of all databases throughout the organization.

Table 5-1 compares the responsibilities of a DA and DBA:

TABLE 5-1

Comparison of Responsibilities of
Data Administration and Database Administration

	Data Administration	Database Administration
Primary responsibility	Administrative	Technical
Scope	All databases	Database specific
Data design	Logical	Physical
Primary liaison	Management	Programmers, analysts
Range of concern	Long-term data planning	More concerned with short-term development and use of databases
Primary orientation	Metadata Data dictionary Data analysis DBMS independent	Data Database Database management systems specific

[Ref. 10: p. 6]

B. BENEFITS OF DATA ADMINISTRATION

The benefits of data administration can be summarized as follows: [Ref. 10: pp. 9-17]

1) Lower costs---The long-term costs associated with data structure and system development are much lower when a

comprehensive data dictionary is used. All future application costs are minor compared to the high initial cost of developing the data dictionary to be shared across several applications.

2) Increased data sharing---Because the data dictionary is comprehensive it allows several applications to share data. As mentioned above this lowers costs significantly in the long-term.

3) Decreased data redundancy---All the planning and logical design work that goes into the data dictionary insures that there is very little data duplication or redundancy. Data modeling, data normalization, and data standards are some of the techniques used by the DA to prevent duplicate data entities.

4) Centralized control and management of data definitions---The DA should be the central repository and control mechanism for all data definitions used by the application development and system maintenance staff. All additions to, changes of, and deletions from data definitions used by application programs and DBMS's should be managed by the DA. This management includes the security, backup, recovery, and audit trail of all changes to data definitions. By centralizing the control of this information, problems with duplicate or conflicting updatesof data definitions can be minimized.

5) Change control---The DA provides the formal documentation and approval process for all changes to metadata.

6) Source of data-design expertise---One of primary duties of the DA is to train, advise, and assist users in the analysis and design of data structures. These data structures include parameter tables, files, databases, records, and segments.

7) Coordination of data usage---The DA is responsible for planning and designing data that will be used for many applications or databases. The DA provides the knowledge necessary for the effective coordination and sharing of information across organizational, project, or individual database boundaries. This minimizes data redundancy and increases the degree of data sharing among the entire organization.

8) End user awareness---Traditional DP duties are today being assumed by the end-users. Some of the new tools being used by the end-users today include: distributed processing, personal computers, report-writers, and query languages. However, these tools are of limited value unless the end-user has access to the data and metadata. Metadata is compiled and maintained by the DA. One of the most important benefits of data administration is to share this metadata with the user community.

C. DATA ADMINISTRATION STANDARDS

Before imposing a set of standards on the DP personnel and end users, the DA should understand the general philosophy and implication of these standards. A good set of rules to follow concerning standards are:

1) No standard is applicable in every situation. However, the DA must not allow exceptions to become the norm.

2) Management must support and be willing to help enforce standards. If standards are violated, management must assist in assuring that the violations are corrected.

3) Standards must be practical, viable, and workable. Standards must be based upon common sense. The less complicated and cumbersome the standards, the more they will be adhered to. Keep standards simple.

4) Standards must not be absolute; there must be some room for flexibility. While some standards must be strictly adhered to, most standards should not be so rigid that they severely restrict the freedom of the data designer.

5) Standards should not be retroactive. Standards are to control and manage present and future actions---not to undo and redo past actions. In most cases, standards enacted today cannot apply to data design that began several months ago.

6) Standards must be easily enforceable. To achieve this, it must be easy to detect violations in standards.

The more the process of auditing for the compliance of standard can be automated, the more effective will be the standards themselves.

7) Standards must be sold, not dictated. Even if upper management wholeheartedly supports DA standards, the standards must be sold to employees at all levels. The DA must be willing to advertise the standards to all employees and to justify the need for such standards. DA standards demand that programmers and analysts change the way they design data. Any lasting and meaningful change must come from the employees themselves.

8) The details about the standards themselves are not important---the important thing is to have some standards. The DA must be willing to compromise and negotiate the details of the standards to be enacted.

9) Standards should be enacted gradually. Do not attempt to put all DA standards in place at the same time. Once standards are enacted, begin to enforce them, but do it gradually and tactfully. Allow ample time for the non-DA staff to react and adjust to the new standards. The implementation of standards must be an evoluntionary, rather than a revolutionary, process.

10) The most important standard in data administration is the standard of consistency---consistency of data naming, data attributes, data design, and data use. [Ref. 10: pp. 31-32]

Before an organization implements any DA standards, it is important that the DA be able to communicate effectively with non-DA personnel. To do this, non-DA personnel must be introduced to some basic data administration/data dictionary terminology. A good way to do this is for the DA to develop a complete glossary of DA terminology and distribute it to all the end users in the organization. [Ref. 10: p. 32]

The standards a DA has to concern himself with are: data element naming standards, standard abbreviations, and a standard way of defining data elements. The following rules are used by the DA and end users to achieve standardized data elements:

1) Define a data element in such a way that the definition of this entity can be adequately described in a single simple sentence.

2) Whenever possible, use combinations or concatenations of generic data elements to identify specific entities.

3) Develop and use standardized and consistent attributes to qualify or categorize data entities.

4) Minimize the use of specific data element names and maximize the use of roles or domain to specify exactness. Strive for modularity in naming data elements much as you strive for modularity in writing programs.

5) When labeling data elements, maximize logical and minimize physical constructs. [Ref. 10: pp. 49-54]

The dictionary name assigned to a data element should be derived from the definition of the data element itself. The dictionary name of an element should reflect the purpose of the entity, not how the element is perceived or used by any one group within the enterprise. A data element should be designed:

---According to logical, not physical, characteristics

---Independent of the hardware or software where it is used.

---Independent of any particular user organization

A data element name should be:

---As meaningful as possible

---Self-documenting

---Easily distinguishable from other data elements in a dictionary

---Derived from the definition of the entity

---A general or generic name

Every data element should be composed of at least:

---One class word

---One prime word

---One or more modifying words

Example: ACCOUNTS-PAY-VENDOR-NUMBER

Class word ---> NUMBER

Prime word ---> VENDOR

Modifier word ---> ACCOUNTS

Modifier word ---> PAY [Ref. 10: pp. 40-41]

D. DATA ADMINISTRATION CHARTER

Because data administration is so important to the data processing function within an organization, it is important to document the objectives and scope of the DA. This document should be reviewed and approved by both the data processing staff and end-user management. The purpose of the data administration charter is to identify the types of authority that the DA requires to effectively perform his administrative duties in managing the corporation's data and system resources. Since data dictionary systems are essential to this task, the DA charter is in part an implementation and usage plan for the data dictionary. [Ref. 15: p. 171]

The DA charter should provide answers to the following questions:

1) How will the DA staff be organized?

2) What are the job descriptions of the members of the DA staff?

3) What level of expertise is needed for DA staff members?

4) What are the accountabilities and responsibilities of the DA?

5) What will be the relationship between the DA and data processing organizations?

6) What will be the relationship between the DA and end-user departments?

7) What is the relationship between the DA and database administration?

8) What are the short- and long-range goals of the DA?

9) Who will be responsible for establishing and maintaining the data dictionaries? [Ref. 10: pp. 6-7]

The DA charter consists of the following three basic sections: objectives, premises, and responsibilities.

Objectives---In this section the basic goals of the DA function are defined. Implicit in such a definition is the organization's perspective on the role of the DA within the corporation, as well as the purpose of general data resource development. The objectives outlined will provide a yardstick for measuring the success of both the DA and his primary tool, the data dictionary. Some examples of typical DA objectives are:

---To improve the quality, authenticity, and timeliness of system documentation.

---To make information about the corporation's data and system resources more available to users both within and outside the DP department.

---To facilitate a migration to data base. [Ref. 15: pp. 185-186]

Premises---This section states assumptions about the organizational position of the DA, optimum strategies for

data dictionary implementation, and the role of the dictionary in data and system management. These assumptions are based on an informed assessment of dictionary technology and of the organizational context in which it is to be used. The premises represent an explicit statement of the thinking on which detailed DA responsibilities will be built. They communicate to management what the range and scope of data dictionary impact will be. Premises collectively define the scope of the DA's responsibilities and thus constitute the heart of the charter as a political statement. It is imperative that the premises be concisely stated and understood and accepted by management before any actual dictionary actions are taken. Some examples of premise statements are:

---Organizational placement of the DA

---DA group staffing

---A corporate language for data [Ref. 15: p. 187]

Responsibilities---In this section, the roles of the data dictionary and of the DA in its management are examined in summary fashion. The framework of new procedures in a variety of areas is examined. The distribution of responsibility between the DA and various classes of dictionary users is detailed for individual aspects of the dictionary's content. The means by which the integrity of the dictionary is to be maintained is stated succinctly. Prior to writing the charter the DA will examine in detail

the following procedural areas and responsibility centers:

1) Application system documentation---Who monitors compliance with departmental standards? How effective are documentation turnover and update strategies?

2) Production systems---What are the procedural steps for giving a new system operational status? For making modifications to an existing system? What information is needed by operations?

3) Naming authority---Who assigns names to job streams, programs, systems, files, reports and databases? Where are such assignments recorded?

4) Copy libraries---What source and object libraries are there currently? How do they fit in with current system implementation methods?

5) Operational information---What information does operations currently keep about production systems? How is it recorded and accessed?

6) System design---What are the approval points for stages in system design? What type of information is required at each point?

7) System implementation---Where are source and object components of test systems kept? Who has responsibility for changes and the communication to affected groups?

8) Reports, data elements, codes---Does the DP department, or any other group in the organization, have special approval cycles for any of these components?

9) Database---What procedural steps are involved in schema and subschema compilation? Where is the source for schemas and subschemas kept? What approval and steps are necessary in modification?

10) Data dictionary---How is responsibility for data entry to the dictionary apportioned among users? What naming and documentation standards exist, and how is compliance monitored and enforced? [Ref. 11: pp. 188-189]

The DA charter is important because it formalizes the DA's position and responsibilities in the organization and documents upper level management support of DA objectives. The charter requires a great deal of planning and cooperation but in the long-run that effort is worthwhile.

Table 5-2 presents final summary of the actions that will determine the success or failure of data administration within an organization:

TABLE 5-2

Keys to DA Success or Failure

Activity	*Do*	*Don't*
1. Plan	Plan short- and long-range goals for DA. Plan how you are going to use the dictionary. Plan the activities within DA to support the business goals of the enterprise. Involve top management in the development and review of these plans.	Don't approach DA or the use of a data dictionary hastily or blindly. Don't begin any DA project without thorough planning.

2. Document	Develop a DA charter and job responsibilities for each job within DA. Put in writing the estimated costs and benefits of all DA efforts before starting. Document DA standards and procedures. Solicit management and user participation and review of these documents.	Don't assume others understand the goals or direction of DA. Don't assume management understands the objectives and limitations of DA.
3. Automate	Automate the population of the dictionary. Automate the auditing for redundancy and compliance with naming conventions. Automate the generation of software from the dictionary.	Don't do any more data dictionary data entry than necessary. Don't manually check for adherence to DA standards. Don't code data definitions manually.
4. Market	Advertise, promote, publicize and sell the benefits of DA and the data dictionary. Invest in education and training for DA principles to the data processing and the end users staff. Devote some time to public relations with the groups that will interface with DA.	Don't dictate or force DA standards. Don't issue commands or edicts concerning DA policies and procedures. Don't expect immediate and complete compliance to new standards.
5. Adapt	Make your standards and data dictionary procedures mesh with the existing environment. Tie DA standards to existing application development guidelines and procedures.	Don't expect the business requirements or company policy to adapt to your rules. Remember, with or without DA, the company must continue to prosper. Don't insist on rigorous controls and compliance before you can support application requirements.
6. Commit	Gain the commitment and support of upper management. Dedicate yourself and others to the successful implementation of DA.	Don't implement on a part time or haphazard approach. Don't underestimate the resources or the time span required to successfully develop the DA function within a company.

VI. A PROPOSED DATA ADMINISTRATION STRATEGY FOR THE U.S. COAST GUARD

A. DEVELOP CORPORATE DATA DICTIONARY

The Coast Guard would benefit from having one central data dictionary. The Coast Guard's information needs could be met with a centralized system as described in Durell's book on data administration [Ref. 10: pp. 134-135]. There are enough data entities common to all applications within the CG to justify a move in this direction. By following the guidelines established in Chapter IV of this thesis a corporate data dictionary could be developed and maintained at CG Headquarters. The twelve CG districts and major field units could have access to this corporate dictionary via transmission through the Standard Terminal (C3) network. Much of the work involved in setting up the corporate dictionary has already been done. [Ref. 3]

Many of the districts have already set up data dictionaries for various data processing applications. Most of these dictionaries have been done according to current relational database theory and concepts including ReQuest^{tm}. What is required now is a thorough review and selection of those data elements that G-T determined could be shared across all CG applications. This would be a significant

first step towards establishing a standard model for data within the CG organization. [Ref. 10: pp. 49-55]

What I am proposing would be a monumental task if G-T attempted to develop the entire CG corporate dictionary at one time. A much better strategy would be to develop the CG corporate dictionary incrementally over several years. This way G-T could solicit help from talented people in the field and continuously add elements to the dictionary after reviewing them for compliance with CG data standards (yet to be developed). This phased data dictionary system acquisition strategy is thoroughly presented by Ross [Ref. 15: pp. 128-168] and Durell [Ref. 10: pp. 31-32].

B. DEVELOP DATA ADMINISTRATION CHARTER/STANDARDS

It is important for an organization to formally recognize its commitment to DA. The best way to do this is to write a DA charter. [Ref. 14: pp. 171-218] The CG should should write a DA charter as a first step towards its commitment to DA.

Of course, writing the charter alone is not enough to guarantee success. DA will succeed only if management and all the end-users are willing to follow the standards set up by the DA staff. The DA staff, on the other hand, has the responsibility of not violating the trust placed in them by upper management and the end-users. They must carefully plan and test every standard before applying it to the the

end-users. Fortunately for DA's there are well-documented and workable techniques for data modeling and data design. A critical factor for many DA's will not be "what" information they present but "how" they represent it. [Ref. 10: pp. 171-175]

C. CONTINUE SOFTWARE EVALUATION BOARD

The Software Evaluation Board (SEB) in its present form should be continued. The voluntary participation by interested personnel from the districts encourages input from people who care about CG data processing issues. Restricting the board to nine members is a good policy because it discourages the members from forming special interest groups. If CG Headquarters and the districts establish DA positions within their respective organizations it would be beneficial to the Coast Guard to include these people on future SEB's. Even if the board continues to limit itself to nine members the input received from the district DA's will surely be well respected and well heeded. The district DA's could become participating but non-voting members of future SEB's.

D. DEVELOP IN-HOUSE ReQuest^{tm} USER TRAINING

As I mentioned earlier it is important that the DA train the end-users in the basic principles of data administration. End-users should be made aware of why data

dictionaries and data standards are so important. Gone are the days when the end-user could rely soley on the DP staff to accomplish everything related to data processing. The end-users have to become familiar with data standards and the data dictionary. They are the ones who will be creating the logical and physical data models. The DA's primary function is to advise the end-users. [Ref. 16: p. ID/36]

Since ReQuest^{tm} was recently chosen as a recommended DBMS for end-users throughout the Coast Guard it would beneficial for CG Headquarters to develop user training for ReQuest as soon as possible. The 13th district's DBMS report [Ref. 7] estimated the average end-user would need 10 days to develop an application using ReQuest^{tm}. Unfortunately, the vendor's estimate of 10 days may be a little on the low side. A more realistic estimate might go be as high as 30-60 days. Another factor to consider is that most end-users don't have ten working days, in one block of time, to devote to learning ReQuest. The end-users more typically spend 1-2 hours a day spread out over 40-80 working days learning ReQuest. Perhaps a more effective way to train CG personnel would be for CG Headquarters to develop a user training program to supplement the training provided by the ReQuest vendor. This in-house training should include: general database/data administration theory and ReQuest^{tm}-specific training.

The Coast Guard should use the "adaptive" training method if it sets up an in-house ReQuest training program. The adaptive technique helps trainees in adapting or applying new technology on an on-going basis. Within each group of trainees, one or more students emerge as natural leaders and teachers. The user/trainers have an aptitude for what they have learned and an ability to pass on their understanding and enthusiasm. The user/trainers should be nurtured by keeping them informed of all matters related to information systems in general and ReQuest^{tm} specifically. The end users will seek out the user/trainers because they are accessible, understand their needs and have the same problems to solve. The user/trainers should also be kept up-to-date on new training materials and new applications. The user/trainers' use of the new technology will often be the most adaptive and should be shared with others in the organization. [Ref. 16: p. ID/36]

The most effective end-user training for ReQuest^{tm} will have the following characteristics:

1) It is targeted to meet specific CG and end-user needs.

2) It is tied to the CG's way of doing business.

3) It uses actual cases or addresses actual problems familiar to the end-users.

4) It consists of ongoing training, with frequent fine-tuning to ensure it meets end-user needs.

5) It uses managers and peers as trainers to promote on-going application of new technology. [Ref. 16: p. ID/36]

VII. CONCLUSIONS

A. REMAIN FLEXIBLE

The DBMS strategy the Coast Guard ultimately settles on should in every case be "flexible". Flexible in the sense that it can adapt to its environment. The environment for data processing is characterized today by technology that is advancing almost exponentially. The potential benefits to be gained from remaining flexible are significant: lower hardware/software costs, lower maintenance costs, greater capabilities, higher productivity, higher DP return on investment, and a higher level of end-user satisfaction. [Ref. 17: pp. 7-18]

A change that is taking place in the current DP environment is the conversion of mainframe DBMS's to microcomputer DBMS's. Microcomputer DBMS sales currently account for less than 10% of total DBMS sales however, it is projected that by 1995 this figure will increase to 33%. [Ref. 20: p. 28] In this highly competitive market the end-user will reap the benefits of these powerful DBMS's in terms of both lower prices and more capabilities.

An example of a powerful new micro DBMS is Cornerstone^{tm} sold by Infocom, Inc. Cornerstone is a DBMS targeted at

managers, small business owners, and other professionals without programming experience, but with personal computer experience. The end-user builds a database by answering a series of simple questions. If the user is unsure of an answer, the package explains the options. The Help system extracts data from the end-user's database and incorporates it into Help messages. After a database has been built, it can be added to or changed without complex system commands [Ref. 18: p. 36]. There may be many DBMS's like Cornerstone appearing in the market soon. The trend in software is currently focused on the end-user. The Coast Guard's overall DP strategy should be adaptable enough to take advantage of these new, powerful micro DBMS's. The CG should not commit itself 100% to any one technology. The CG would benefit more by setting aside money for purchasing new technologies or investing in research and development that would ultimately result in new technologies.

B. FOCUS ON THE END-USER

Just as the DP market is now catering to the end-user the Coast Guard also needs to recognize the end-user as the most important factor in its DA strategy. There are several techniques to accomplish this goal. One is the end-user committee. The end-user committee is a team of users who have expert knowledge about their own data. They meet periodically with the data administrator who works with them

to design data structures which are then input into the data dictionary and logical data model. [Ref. 19: p. 193]

Another very effective way to assist the end-user is to establish an information center. The 13th Coast Guard District has implemented such a system and it appears to be working quite successfully [Ref. 5]. They support their end-users by training them how to use the hardware and software available to them. This is a powerful concept and one that should be adopted by all districts in the Coast Guard.

C. PLAN FOR DATA ADMINISTRATION

The current Coast Guard DP environment includes a commitment on the part of upper level management to implement current DBMS technology. In this dynamic environment I believe it is extremely important that the Coast Guard begin directing its efforts towards "data" planning and management vice "process" planning and management (which includes both hardware and software resources). Data is a resource the Coast Guard must develop and protect if it is to take full advantage of current and future DP technologies. The data administration concepts presented in Chapters IV, V, and VI of this thesis will hopefully give some guidance to implement data administration in the Coast Guard.

The Coast Guard has the knowledgeable end-users needed to define data elements for a corporate data dictionary. What is needed now is a carefully planned and coordinated effort to apply standard names to the data objects and to develop a common logical data model for the Coast Guard organization.

Once a corporate data dictionary has been implemented the Coast Guard will be able to share data across many applications and possibly access that data via many DBMS's. In the long run the hardware and software resources of the Coast Guard will undoubtedly change many times. However, the data will remain relatively unchanged. Standardizing this data and protecting it will improve operations and reduce costs for the Coast Guard in the long run.

LIST OF REFERENCES

1. Nolan, Richard L., "Managing the Crisis in Data Processing," Harvard Business Review, March-April 1979.

2. USCG, Automated Data Processing (ADP) Plan, COMDINST M5230.8B, 7 Dec 1983.

3. USCG, Standard Terminal Management Plan, Enclosure (1) to COMDTINST M30909.2, 24 Aug 1983.

4. Squires, Fred N., LCDR, USCG, G-T Report: Managing End-User Conputing --- One Agency's Approach, 28 May 1984.

5. USCG, 13th District Instruction 5200.1, Information Center Training and Consultant Services, 28 Oct 1983.

6. USCG, 12th District Report, Aids To Navigation Information System (ATONIS), 1 Oct 1984.

7. USCG, 13th District Report, Electronics Data Systems Corporation Report --- Database Management Systems Evaluation, 18 Sep 1984.

8. "Database Program's Popularity Soars Following Debut," Computerworld, 24 Sep 1984.

9. "Contract Awards," Government Computer News, Vol. 2, Sep 1983.

10. Durell, William R., Data Administration ---A Practical Guide to Successful Data Management, McGraw-Hill Book Co., 1985.

11. Weldon, Jay-Louise, Data Base Administration, Plenum Press, 1981.

12. Coulson, Christopher J., "Six Steps to Successful Data Dictionary Implementation - Part 1," Database Newsletter, Vol. 11 No. 2, Mar/Apr 1983.

13. Connall, Duncan G., "Choosing Standard Dictionary Names," Database Newsletter, Vol. 12 No. 2, Mar/Apr 1984.

14. Coulson, Christopher J., "Six Steps to Successful Data Dictionary Implementation - Part 2, Database Newsletter, Vol. 11 No. 3, May/Jun 1983.

15. Ross, Ronald G., Data Dictionaries and Data Administration --- Concepts and Practices for Data Resources Management, AMACOM, 1981.

16. Kerns, Steven, "What End-Users Don't Know May Hurt You," Computerworld, 8 Oct 1984.

17. Martin, James, An End-User's Guide to Data Base, Prentice-Hall, Inc., 1981.

18. "Cornerstone, Database for Nonprogrammers," Software News, Vol.4 No. 12, Dec 1984.

19. Martin, James, Managing the Database Environment, Prentice-Hall, Inc., 1983.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, CA 93943	2
3.	Department Chairman, Code 54 Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93943	1
4.	Professor Norman R. Lyons Code 54Lb Naval Postgraduate School Monterey, CA 93943	1
5.	LT Barry Frew Code 54Fw Naval Postgraduate School Monterey, CA 93943	1
6.	Commandant (G-TPP) ATT: Mr. Russ Cherry United States Coast Guard 2100 Second St., S.W. Washington, DC 20593	1
7.	Commandant (G-TPP) ATT: Mr. Howard Pritchard Jr. United States Coast Guard 2100 Second St., S.W. Washington, DC 20593	1
8.	Commandant (G-PTE) United States Coast Guard 2100 Second St., S.W. Washington, DC 20593	2
9.	LTJG Thomas A. Jarrad 11th Coast Guard District (dt) 400 Oceangate Long Beach, CA 90822	1

10. LT Mark L. Scire 1
1043 Fifth St., Apt #1
Monterey, CA 93940

211917

Thesis
J2944 Jarrad
c.1 Proposed Data Administration strategy for the U.S. Coast Guard.

211917

Thesis
J2944 Jarrad
c.1 Proposed Data Administration strategy for the U.S. Coast Guard.

thesJ2944
Proposed Data Administration strategy fo
3 2768 002 10030 7
DUDLEY KNOX LIBRARY